R E P O R T R E S U M E S

ED 011 338 24

PROBLEMS OF MEASURING SPEECH RATE.

BY- CARROLL, JOHN B.

REPORT NUMBER BR-5-0215-38 PUB DATE 19 OCT 66

CONTRACT OEC-5-10-239

EDRS PRICE MF-$0.09 HC-$0.52 13P.

DESCRIPTORS- *READING SPEED, ORAL COMMUNICATION, *ORAL READING, *MEASUREMENT TECHNIQUES, METHODOLOGY, VERBAL COMMUNICATION, *SPEECH, *SPEECH COMPRESSION, LOUISVILLE, CAMBRIDGE

A DISCUSSION WAS PRESENTED ON THE PROBLEMS OF MEASURING SPEECH RATE, A CRITICAL VARIABLE IN SPEECH COMPRESSION, BOTH IN DESCRIBING THE INPUT TO ANY SPEECH COMPRESSION SYSTEM AND IN CHARACTERIZING THE OUTPUT. THE DISCUSSION WAS LIMITED TO SPEECH RATE MEASUREMENT OF "ORAL READING RATE," ONLY, AND DID NOT DEAL WITH THE MEASUREMENT OF "SPONTANEOUS SPEECH." IT WAS REPORTED THAT IF MEASUREMENTS ARE TAKEN AS "UNITS OF SPEECH OUTPUT PER UNIT OF TIME," THEY WILL NORMALLY BE DISTRIBUTED OVER PERSONS OR OCCASIONS, BUT MEASUREMENTS WILL BE POSITIVELY SKEWED IF TAKEN AS "AMOUNT OF TIME PER UNIT OF PERFORMANCE." THE USE OF "UNITS OF TIME PER UNIT OF SPEECH OUTPUT" WAS SHOWN TO BE MORE EFFECTIVE IF VERY SMALL TIME UNITS ARE USED TO REPORT SPEECH RATE DATA. "WORDS PER MINUTE" WAS DISCOUNTED AS A UNIT OF SPEECH BECAUSE WORDS VARY IN LENGTH. AS PHONEMES ARE DIFFICULT TO COUNT, THE SYLLABLE WAS RECOMMENDED AS THE UNIT OF SPEECH OUTPUT FOR MEASURING SPEECH RATE, WITHOUT ACCEPTING REPRESENTATIVE VALUES OF THE UNITS. THESE VALUES COULD BE STANDARDIZED BY FURTHER RESEARCH. THE AUTHOR CONCLUDED WITH A DISCUSSION ON THE COLLECTION OF NORMATIVE DATA FOR STANDARDIZING THE USE OF SPEECH UNITS FOR PRECISE ORAL READING MEASUREMENTS. THIS PAPER WAS PREPARED FOR PRESENTATION AT A CONFERENCE ON SPEECH COMPRESSION, UNIVERSITY OF LOUISVILLE, OCTOBER 19, 1966. (JH)

ED011338

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

BR-5-0215-38
OEC-5-10-239
PA 24

Problems of Measuring Speech Rate

John B. Carroll
Harvard University

[For presentation at a conference on Speech Compression, U. of Louisville, Louisville, Kentucky, October 19, 1966]

Speech rate is clearly a critical variable in speech compression, both in describing the input to any speech compression system and in characterizing the output. Reliable and meaningful measures of speech rate are "musts" if we hope to appraise what a speech compression system actually does. It cannot be assumed that just because a system has a certain compression ratio, such as 75% , it will always produce speech at a certain rate, because obviously the output rate will depend upon input rate as well as the compression ratio. This fact seems to have been forgotten by some of the more ardent exponents of speech compression when they report output rates of, say, 500 wpm by merely using the compression ratio in their calculation, without making careful measurements of the input rates.

As I will show, measuring speech rate is not a simple matter of counting words per minute, for this measure can give misleading results.

Measurements of speech rate enter our considerations at another point. A tacit assumption in most speech compression work is that this process will have an output that is in some way unusual, i.e. not normally producible by the unaided speaker. This raises the question of what is in fact normal, usual, or typical, and also the question of where the dividing line is between what is normal, usual, and typical and what is abnormal, unusual, or atypical. To answer this question we need not only reliable normative data on normal speech rates but also psychophysical studies of subjective responses to speech at various rates (Hutton, 1954). Although the literature contains some leads on these matters, it is not adequate for our purposes.

Further, we need to know something about the ability of speakers to control their speech rates at designated levels or to vary them on demand. This is undoubtedly an important matter in connection with the production of speech recordings for the blind. Likewise we need to know much more about the ability of listeners to comprehend materials at various rates; in research on this matter the different procedures in the measurement of speech rate must be taken carefully into account in interpreting the results.

In this talk, I will consider various problems of measuring speech rate as they pertain to what is ordinarily called oral reading rate, that is, the rate at which a speaker reads aloud a continuous prose text which he has not necessarily seen before. This is in contradistinction to spontaneous speech rate, i.e. the rate of speaking when the speaker is continuously composing speech of a "novel" character. I shall not deal with spontaneous speech rate, which entails some very difficult problems of measurement, although some of the same considerations apply to it as in the case of oral reading rate.

Let us suppose that our problem is to determine the distribution of the oral reading rates of a large sample of educated adult speakers of English; if the problem is stated in this way it entails practically all the problems I want to discuss and excludes such problems as the sampling of speakers, with which I am not concerned. I should note, however, that in reporting norms for reading rates, the sampling of the speakers must be adequately specified in terms of age, sex, education, intelligence, background, and any abnormal characteristics such as speech handicaps, deafness, etc.

First, we must know what we mean by rate. In the present context, rate must be reported according to the following model:

X units of speech output per unit of time.

For example, we might report words per minute, syllables per second, or even pages per hour, if the pages are of some specified standard size and wordage.

We shall speak of the problem of units of measurement in a moment. Right now, I call your attention to the fact that the model given as

X units of speech output per unit of time

is not the same as

Y units of time per unit of speech output.

This second model is exemplified by such measurements as minutes per 300 words, seconds per syllables, or minutes per page. To be sure, these two models are related to one another; by doing the appropriate arithmetic, one can be obtained from the other. However, reporting a rate in terms of the second model does not give measurements that can be as readily apprehended and handled statistically as by the first model. 300 syllables per minute is a more readily grasped measurement than the equivalent statements ".33 minutes per 100 syllables" or "0.2 seconds per syllable," Furthermore, taking arithmetic averages of measurements reported in output per unit of time is legitimate, while taking arithmetic averages of measurements reported in time per unit of output is not. In fact, the two averaging procedures give different results; if the second kind of measurement is to be averaged at all, the harmonic mean rather than the arithmetic mean must be used.

All this is important because a surprising number of investigations have used improper procedures. In the literature, the most common procedure for measuring speech rate is to clock a speaker while he reads a standard paragraph. Thus, Johnson (1961) clocked the times (in seconds) to read a 300-word passage. There is nothing essentially wrong in this, but when he proceeded to average these times, using the arithmetic mean, the result was misleading since it was not the same as the average he would have gotten by converting the times to the form Words per minute. Furthermore, generally speaking, measurements taken as units of output per unit of time are normally distributed over persons or over occasions, while measurements taken as amount of time per unit of performance are skewed positively.

If very small time units are used, data must be reported to sufficient precision to permit projections to larger units. For example, a measurement reported as "3 syllables per second" is probably imprecise, unless it is actually "3.000 syllables per second," from which one could reliably translate to "180.0 syllables per minute."

The real bugaboo in this picture is the question of units. Units of time, at least, can be specified without difficulty and can be measured to almost any desired degree of precision if sufficient precautions are taken, although even here, errors can be made. I remember an occasion when one of my assistants timed something with a stopwatch graduated in 100ths of a minute, under the impression that it was graduated in the normal seconds and fifths of a second. Naturally the results were disastrous. But when we consider units of speech output, there are some real problems. It is apparently a standard convention to report speech rates in terms of words per minute, and thus far in this talk I have spoken about words per minute just in order to seem conventional. From the standpoint of scientific measurement, however, the word is a very poor unit and I wish that we could abolish wpm measurements. It is not a standard unit, for words vary in length from the simple a to the classic antidisestablishment-arianism. Furthermore, different samples of prose vary considerably in the average length of their words, whether measured in terms of syllables or in terms of phonemes. Number of syllables per word figures as an element in various measures of readability.

Let me report in this connection an experiment that Jeffrey Sampson did under my supervision at Michigan recently. Six passages from American fiction were selected, each containing between 300 and 320 words and from 10 to 15 sentences. In most respects, the passages were homogeneous both within and among themselves; the one respect in which they were deliberately allowed to vary

was the syllable-to-word ratio (S/W). These ranged from 1.20, for a passage from Ernest Hemingway, to 1.73, for a passage from Henry James. Twenty-four speakers were asked to read these passages aloud "at a normal, comfortable rate" and they were timed. The order in which the passages were read was systematically varied according to a Latin square; the Ss were not aware that the passages varied in S/W ratio. The times for individual subjects were converted into words, syllables or phonemes per minute and then averaged. Table 1 and Figure 1 show the results: It is evident that words/min is systematically related to S/W; syllables/min much less so, and phonemes/min still less so. Furthermore, the coefficient of variation of the means over passages was greatest in the case of words and least in the case of phonemes; in other words, phonemes/min. gave the most consistent results. Syllables do, of course, vary in length, but average length of syllable probably does not vary from text to text as much as word length. (For the six passages, phonemes/syllable ranged from 2.71 to 2.54 and tended to be inversely correlated with S/W ratio; evidently the syllables of longer words tend to be shorter than the syllables of short words.) Nevertheless, since phonemes are difficult to count, for practical purposes I recommend the syllable as the unit of speech output in measuring speech rate.

But because the average speech rate in terms of even syllables/minute varies somewhat with the nature of the text, it is probably wise to collect norms on a rather wide variety of texts, before we can accept really representative values.

There is an interesting sidelight from this study; namely that oral reading rates differed very consistently over the 24 speakers. Correlations of rates among the six passages ranged from .91 to .98. This means that each speaker maintained a very consistent rate over the six passages, and these rates varied widely.

A number of investigators (e.g. Starkweather, 1960; Shearn et al., 1961) have used apparatus-produced units; i.e., they report the rate at which the intensity of the speech signal passes specified levels determined by the settings of electronic gating circuits. Two things should be recognized about these measurements: one, they may be affected by the nature of the texts read, and two, they are certainly affected by the parameters of the electronic circuits. Such measurements should always be accompanied by equations whereby the results can be converted to more standard units such as syllables.

Let me speak more about the collection of normative data. The rate at which a speaker will read a text is obviously a function of his set. We do not know to what extent a speaker can consciously control his oral reading rate, but in collecting norms we can distinguish at least three distinct sets that can be controlled by instructions. One: we can ask a speaker to read "as fast as he can"; Two: we can ask a speaker to read "at a normal, comfortable rate." (This was the instruction used in the data collected by Sampson.) Three: we can ask a speaker to read "so as to communicate"---"as if he were reading the material to a friend, making sure that he gets the meaning." Schwartz (1961) demonstrated that speakers instructed to "read so as to communicate" slowed down considerably from the "normal" rate; from his data I compute that the rate was on the average about 87% of the normal rate. (Schwartz, incidentally, failed to report rates in proper form. He reported his measurements in terms of average number of minutes to read 5 sentences; thus, he committed three errors: (1) failure to specify units of speed output in useful terms; (2) reporting rate as time per unit of speech output; and (3) arithmetic averaging of times per unit of output.)

There are several instances in the literature in which speakers were reported to control their rates at designated levels, but one gets the impression that this was accomplished only after considerable trial and error. Harwood's

speaker (1955) read at "carefully controlled" rates of 125, 150, 175, and 200 words per minute; Goldstein"s speaker (1940) read at various rates from 100 to a surprising 322 wpm. Note, of course, the unfortunate use of words as the units in these measurements.

I can report a bit of normative data using the "maximum" speed set and the "normal" speed set. My subjects were 130 college and graduate students, all native speakers of English. For a text of 125 words or 218 syllables (thus, syllable-to-word ratio = 1.74), the rates under the "maximum" speed set were normally distributed with the following means and standard deviations:

	Mean	S. D.
WPM	205.4	29.6
SPM	358.2	51.6

These data can be compared with the data for passage 6 in Figure 1, since the syllable/word ratio was similar. For a text of 177 words or 297 syllables (thus, syllable-to-word ratio = 1.68), the rates under the "normal" speed set were normally distributed with the following means and standard deviations:

	Mean	S. D.
WPM	172.2	19.2
SPM	289.2	32.2

The latter data are consistent with the data reported in Figure 1 if account is taken of the S/W ratio.

The literature contains a number of reports of oral reading rates (Darley, 1940; Gibbons et al., 1958; Peters, 1954; Henze, 1953; Starkweather and Hargreaves, 1964), but in most cases the reports are difficult to interpret because they are only in terms of words per minute or because insufficient information is given as to the instructions to the speakers, the nature of the texts read, and other relevant variables.

I will conclude by saying that in this brief span of time I have been unable to cover the literature available or to mention all the variables that appear to be relevant. I have, however, tried to identify the main issues that must be reckoned with in measuring speech rates, whether they are inputs or outputs of speech compression systems.

REFERENCES

Darley, Frederic L. A normative study of oral reading rate. M. A. Thesis, State University of Iowa, 1940.

Gibbons, E. W., Winchester, R. A.., and Krebs, D. F. The variability of oral reading rate. J. speech and hearing Disorders. 1958, 23, 591-593.

Goldstein, Harry. Reading and listening comprehension at various controlled rates. Teachers College Contributions to Education, No. 821. New York: Bur. Publ., Teachers College, Columbia Univ., 1940.

Harwood, Kenneth A. Listenability and rate of presentation. Speech Monographs, 1955, 22, 57-59.

Henze, Reinhold. Experimentelle Untersuchungen zur Phänomenologie der Sprechgeschwindigkeit. Z. exp. angewand. Psychol., 1953, 1, 214-243.

Hutton, Charles L., Jr. A psychophysical study of speech rate. Dissertation Abstr., 1955, 15, 168. (Unpub. Ph.D. Thesis, Univ. Iowa, 1954)

Johnson, Wendell. Measurements of oral reading and speaking rate and disfluency of adult male and female stutterers and nonstutterers. J. speech and hearing Disorders, Monogr. Suppl., 1961, No. 7, 1-20.

Peters, Robert W. The effect of changes in side-tone delay and level upon rate of oral reading of normal speakers. J. speech and hearing Disorders, 1954, 19, 483-490.

Schwartz, M. F. Differential effect of instructions upon the rate of oral reading. J. acoust. Soc. Amer., 1961, 33, 1801-1802.

Shearn, Donald, Sprague, R., and Rosenzweig, S. A method for the analysis and control of speech rate. J. exp. Anal. Behav., 1961, 4, 197-201.

Starkweather, John A. A speech rate meter for vocal behavior analysis. J. exp. Anal. Behav., 1960, 3, 111-114.

Starkweather, J. A., and Hargreaves, W. A. The influence of sodium pentobarbital on vocal behavior. J. abnorm. soc. Psychol., 1964, 69, 123-126.

Table 1

Data on Oral Reading Rates for Six Passages Varying in Syllable/Word Ratio, by Three Methods of Measurement, from 24 Adult Subjects.

	Passage[1]					
	1	2	3	4	5	6
Words	302	319	305	319	318	313
Syllables	362	411	418	475	512	542
Phonemes	980	1071	1108	1205	1315	1380
Syll./Wd.	1.20	1.29	1.37	1.49	1.61	1.73
Phon./Syll.	2.71	2.61	2.65	2.54	2.56	2.54

Oral Reading Rates								Means over Passages
Words/Min.	M	221.75	224.47	191.76	187.16	174.38	169.85	194.90
	σ	30.71	26.63	24.35	21.92	21.40	21.69	21.22
	$M/\bar{M}$	1.14	1.15	.98	.96	.89	.87	C.V. =
	σ/M	.16	.14	.12	.11	.11	.11	.109
Syll./Min.	M	265.81	289.22	262.80	278.64	280.77	294.12	278.56
	σ	36.82	34.31	38.38	32.60	34.46	37.56	11.34
	$M/\bar{M}$	.95	1.04	.94	1.00	1.01	1.06	C.V. =
	σ/M	.13	.12	.12	.12	.12	.13	.041
Phonemes/Min.	M	719.60	753.64	696.60	706.89	721.09	748.61	724.41
	σ	99.67	89.40	88.49	82.73	88.48	95.64	20.28
	$M/\bar{M}$	.99	1.04	.96	.98	1.00	1.03	C.V. =
	σ/M	.14	.12	.12	.11	.12	.13	.028

[1] The passages were slightly edited versions of selections as follows:

1. Ernest Hemingway, The Old Man and the Sea. (1952)
2. Sherwood Anderson, Winesburg, Ohio. (1919)
3. F. Scott Fitzgerald, Tender is the Night. (1933)
4. Thomas Wolfe, Look Homeward, Angel. (1929)
5. J. D. Salinger, Franny and Zooey. (1955)
6. Henry James, Daisy Miller. (1878)

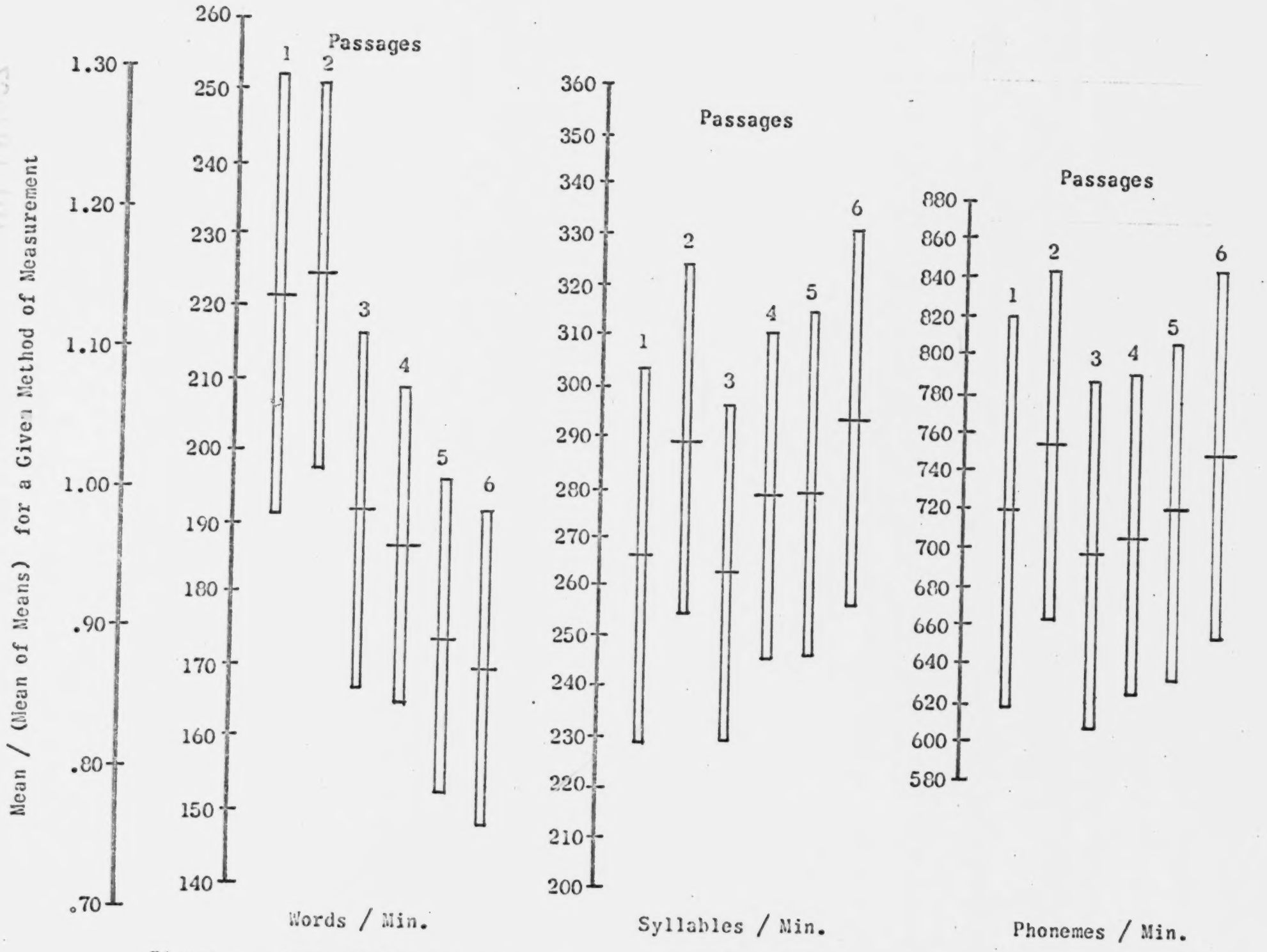

Figure 1. Oral reading rates for six passages varying in syllable/word ratio, by three methods of measurement. Each bar shows the mean and $\pm$ 1 σ for the distribution of reading rates for 21 adult subjects. The vertical scale at the left makes the data comparable over the three methods of measurement, reflecting the absolute amount of variance among the means for a given method.